

- 1 -

"A method and system for quality of service optimisation in a data network"

- 5 The present invention relates to a method of estimating the optimum service rate at a specified quality of service for the transmission of packets of data of different characteristics through a switch node comprising a buffer having a defined size. The invention is also directed towards a system for carrying out such a method.
- 10 Traffic in a data network is essentially composed of individual transactions or flows from a source. The source generates a stream of bits grouped into packets. A complete description of the statistics of the source would be very complex. If the goal is the provisioning of sufficient resources at a router or switch in order to limit packet loss and delay at that router, then a more compact encapsulation of the statistics of traffic arriving at the router can be sufficient. We refer to any such encapsulation as a traffic descriptor. One such traffic descriptor has been described in the US Patent Specification No. 6580691 (Bjoerkman et al), namely a polygonal approximation to a scaled cumulant generating function (sCGF). This US Patent Specification discloses a method and system for estimating the sCGF on-line in real time and storing it as a compact traffic descriptor.
- 15 20 The terminology of this specification is that which is used in High-Performance Communication Networks [Jean Walrand and Pravin Varaiya (Second Edition) Academic Press 2000], unless clearly otherwise.
- 25 30 A communication network is a collection of network elements interconnected so as to support the transfer of information from a user at one network node to a user at another. The principal network elements are links and switches. A link transfers a stream of bits from one end to another at a specified rate with a given bit error rate and a fixed propagation time. In this specification, we refer to the rate in bits per second at which a buffer is served as the service capacity. Other terms often used interchangeable with service capacity are link-rate and bandwidth. Links are unidirectional. The most important links are:
- optical fibre;

- 2 -

- copper twisted pair;
- ethernet.

Several incoming and outgoing links meet at a switch, a device that transfers bits from its incoming links to its outgoing links. The name "switch" is used in telephony, while in computer communications, the device that performs routing is called a router; the terms are used interchangeably in this specification. When the rate of incoming bits exceeds that of outgoing bits, the excess bits are queued in a buffer at the switch. The receiver of each incoming link writes a packet of bits into its input buffer; the transmitter of each outgoing link reads from its output buffer. The switch transports packets from an input buffer to the appropriate output buffer.

The quality of a communications network service, as perceived by a user, varies greatly with the state of the network. To make packet-switched networks economically viable, it is necessary to be able to guarantee quality while reducing capital investment and operating expenses.

Degradation in the perceived quality of a service can often be traced back to loss or delay of data packets at a node or switch in the network. User satisfaction can be guaranteed by managing loss and delay of packets at those nodes where congestion can occur.

Typically, users transmit bits in bursts: active periods are interspersed with periods of inactivity. The peak rate of transmission cannot exceed the link rate. The mean rate of transmission, by definition, cannot exceed the peak rate. The ratio:

$$\frac{(\text{peak rate}) - (\text{mean rate})}{(\text{mean rate})}$$

is a measure of the burstiness of the source.

Loss and delay of data packets at a node in the network arise from the queuing of packets in the buffers of switches or routers. Buffers are required to cope with fluctuations in the bit-rate on incoming links. However, if the buffers are too small,

- 3 -

packets will be lost as a result of buffer overflow; if the buffers are too large, some packets will experience unacceptable delays. For a given buffer-size, loss and delay can be reduced by increasing the capacity of the outgoing link.

- 5 To eliminate packet loss entirely, it would be necessary to increase the capacity of the outgoing link to equal the sum of the capacities of the incoming links. This is prohibitively expensive. Nevertheless, it is a strategy employed sometimes by network operators who take a conservative view on assuring network quality of service.

10

There is a better way. It is unnecessary to eliminate packet loss and unacceptable packet delay in order to give satisfactory perceived quality. It is enough to keep their frequency within predetermined bounds: These bounds are referred to as Quality of Service (QoS) targets.

15

The optimal way to ensure satisfactory perceived quality is to provide the minimum capacity that will guarantee the QoS targets. This minimum capacity is referred to as the Bandwidth Requirement (BWR) of the bit-stream. It lies somewhere between mean rate and the peak-rate requirement.

20

The existence of a BWR and its value can be demonstrated experimentally with a router by observing the change in the frequency with which a target queuing delay in an output buffer is exceeded when the capacity of the outgoing link is varied.

25

The mean-rate and the peak-rate do not depend on the QoS targets. For bursty traffic, the peak-rate can be many multiples of the mean-rate. As the QoS target changes, the BWR varies between them.

30

For a given QoS target, the BWR depends strongly on the nature of the traffic. There is no universal multiplier than can be applied to the mean-rate or peak-rate to give the BWR for a given QoS target. The present applicants have provided various ways of measuring BWR on line in real time such as described in US Patent Specification No. 6580691 (Bjoerkman et al).

- 4 -

This opens the way for many applications: monitoring network quality levels, QoS-sensitive service provisioning, IP call admission control, traffic-based billing and capacity planning.

- 5 Essentially, therefore, given the buffer size b and the QoS target Q , the BWR of the communications system can be calculated from the traffic descriptor D . It will be appreciated that the traffic descriptor, which is essentially the statistical properties of the data, is all important. That descriptor must contain sufficient statistical information to allow computation of the bandwidth requirement. Thus, it is vital to choose the
10 correct Descriptor format, definition or methodology to describe the statistical properties of the traffic. Essentially, the traffic descriptor D describes the characteristics of the particular traffic.

15 There are circumstances in which the nature of the source is such that the statistics of the packet stream are independent of the service capacity available at the router in question. This is the case, for example, where the output of the source is voice traffic – digitised speech. Here, the statistics depend on the behaviour of the speaker and on the codec used to digitise the audio signals, but not on the available service capacity. In general, we refer to packet streams from sources whose statistics are
20 independent of the available service capacity as inelastic traffic.

25 However, there are circumstances in which the nature of the source is such that it cannot be assumed that the statistics of the traffic stream is independent of the service capacity at the router. A feedback mechanism causes the statistics of the source to depend on the service capacity S available to its packet-stream at the buffer. One example of this is web-browsing; here the rate at which the user requests web-pages can depend on the rate at which they are received. In general, we refer to packet streams from sources whose statistics depend on the available service-capacity as elastic traffic.

30

US Patent Specification No. 6,266,322 B1 (AT&T Corp) which is specifically referenced in totality herein describes in some detail how elastic data traffic is handled.

- 5 -

Indeed this US patent specification discloses a method of dimensioning link bandwidth for elastic data traffic for a link in a communications network. However, this US patent does not disclose a method of accurately determining the minimum service rate to maintain quality of service requirements within the network. Further it
5 does not address the problem of selecting an optimum service rate. It could be suggested that it is not as important for elastic traffic as inelastic traffic, since the former operates by feedback control to adapt to time-varying available bandwidth.

PCT Publication No. WO 01/13557 (Fujitsu Network Communications, Inc.) discloses
10 a system for supporting multiple application traffic types over a connection network, for example, elastic and inelastic traffic. Further examples of patents in the area of handling elastic traffic in data networks are disclosed by US Patent No. 6,115,359 (Nortel Networks Corporation) and PCT Publication No. WO 01/28167 (Telefonaktiebolaget LM Ericsson). However, this patent does not teach how to
15 address, in a fundamental way, the problems of inelastic traffic. It relates more to committing some bandwidth to the inelastic traffic by mapping the inelastic traffic to components or switches which provide bandwidth commitments. These commitments enable the data to be transferred through the switch without exceeding a predetermined delay. However, it could be suggested that it is still not accurately
20 estimating the bandwidth requirement. It is more a question of allocating sufficient bandwidth having regard to what it has believed the traffic to be.

The fundamental problem in providing quality of service in a data network is to determine accurately the optimum service rate which is in effect the minimum service
25 rate at which packets of data are removed from a buffer in a switch, whilst maintaining quality of service in the data network and optimising the available bandwidth in the network. Methods devised for determining this optimum service rate for inelastic traffic do not work without modification for elastic traffic because the statistical character of the traffic changes in respect to changes in the available
30 service. In practice, data network operators use trial and error methods in setting the service rate to achieve the desired response time for delivering elastic traffic. This is costly and time consuming as it requires network operators to continually monitor the response time and reconfigure the service rate.

- 6 -

The present invention is directed towards providing a method of estimating the optimum service rate at a specified quality of service for the transmission of data of different characteristics through a switch having a defined size. The invention is
5 directed towards doing this without the need for human intervention.

Statements of Invention

According to the invention, there is provided a method of estimating the optimum
10 service capacity at a specified quality of service (QoS) for the transmission of packets
of data traffic of different characteristics, the traffic being described by a
predetermined type of descriptor (D) to allow the calculation of the estimated
bandwidth requirement (BWR) for that traffic, through a switch node comprising a
buffer having a defined size (b), comprising at time intervals, carrying out the steps
15 of:

- (a) configuring the service-capacity;
- (b) sampling the traffic;
- (c) extracting the descriptor (D);
- (d) calculating from the descriptor (D) the BWR for the configured
service capacity;
- (e) using the calculated BWR to configure a new service capacity;
- (f) iteratively carrying out steps (b) through (e) until the calculated
BWR and the configured service capacity coincide to provide a
final service capacity; and
- (g) defining this final service capacity as the optimum service capacity
for that traffic at that buffer.

A considerable advantage of the present invention is that it operates successfully without a knowledge of the precise form of the bandwidth requirement (BWR) for various service rate at a buffer of a switch. Indeed, relatively little knowledge is known. All that is required is to configure a service capacity, calculate the BWR to 5 configure a new service capacity and keep on repeating this until the calculated rate of BWR and configured service capacity coincide. It is relatively simple and easily carried out.

In one way of carrying out the invention, the initial service capacity specified in step 10 (a) is the previous optimum service capacity. This will lead to a quicker calculation as generally, working off a new optimum service capacity is most likely to lead to a service capacity relatively close to the previous optimum service capacity.

Ideally, the traffic is continuously monitored so that if the nature of the traffic changes, 15 a new optimum service capacity is calculated. Further, it is envisaged that when the required target QoS changes from the target QoS initially set, the target QoS is reset and a new optimum service capacity is calculated.

It is envisaged that some or all of the various steps (a) to (g), as listed above, can be 20 carried out in various jurisdictions, other than the step of configuring the service capacity for the transmission. Further, it will be appreciated that the method further comprises using the optimum service capacity to control the transmission of the traffic through the switch node.

25 In another method according to the invention, it is possible that a sample of the traffic is received with step (a), as listed above, being carried outside the jurisdiction and in which at least steps (c) to (g) are carried out within the jurisdiction. Again, it will be appreciated that the switch node may be remotely located with respect to where the method is carried out, except for the downloading of data to and from the switch 30 node.

Further, the invention provides a closed loop control system comprising a communications network in which are interconnected:-

- 8 -

- user end systems for the delivery and reception of data;
- a switch node incorporating at least one buffer;
- 5 means to configure a service rate; and
- a programmable controller having means to carry out the method of any preceding claim.
- 10 With this latter closed loop control system, the controller may be directly connected to a specific end user output source for the transmission of data to the switch node.
- Further, the invention provides a computer program comprising program instructions for causing a computer to perform the method as laid out above.
- 15 In a further embodiment, the computer program comprises program instructions for causing a computer to provide the means as laid out above. Such a computer program may be embodied on a record medium, a computer memory, a read only memory or carried on an electrical signal carrier.
- 20
- Detailed Description of the Invention**
- The invention will be more clearly understood by the following description of an embodiment thereof given by way of example only with reference to the
- 25 accompanying drawings in which:
- Figs. 1 to 4 are various graphs illustrating bandwidth requirement as a function of service rate,
- 30 Fig. 5 is a flowchart illustrating the iterative process to estimate the optimum service rate, and
- Fig. 6 is a hardware configuration of part of a communications system.

- 9 -

As stated above, for resource provisioning, it is important to know the minimum service-capacity that must be provided at a buffer in the router to ensure that a given Quality of Service (QoS) target is met. We shall denote this rate in bit/sec. by BWR and generally refer to it as the required service capacity. The BWR depends
5 on three factors. The first factor is the quality of service (QoS) that you require which may for example in the transmission of elastic data be described as the number of packets dropped i.e. those that have to be resent. In other words they are not accepted at a particular buffer and must be resent by the source. Typically as explained above, 1 in 10^5 packets dropped or alternatively that 1 in 10^3 packets are
10 not delayed by more than 30 milliseconds are often considered acceptable.

The second factor is the size of the buffer which is preconfigured.

15 The third factor that effects the BWR is effectively the nature of the actual data being transmitted. This as already explained, can be described statistically and there are various methods of providing such a descriptor which relate to certain statistical properties of the data being transmitted. Accordingly therefore one can describe the bandwidth requirement as a function:

20 $BWR = f(D, b, QoS)$

where:

D – is the descriptor of the relevant statistical properties of the traffic flow.

b – buffer size

25 QoS – Quality of Service

For elastic traffic, the traffic descriptor is a function of the available service-capacity:

30 $D = g(S)$

The invention describes how an estimator designed to determine BWR for inelastic traffic can be used to determine the minimum service required to ensure that a given QoS target is met for elastic traffic. For a given service-capacity S, the estimator

- 10 -

yields a value $B(S)$ given by

$$B(S) = f(g(S), b, QoS)$$

5 It is observed experimentally that:

- (a) $B(S)$ is an increasing function of S
- (b) For small values of S , $B(S)$ is greater than S , indicating that the QoS target is
10 not achieved when the service-capacity has the value S .
- (c) For large values of S , $B(S)$ is less than S , indicating that S exceeds the value required to achieve the QoS target.

15 It follows that the BWR for elastic traffic is that value S' of the service-capacity for which

$$S^* = B(S^*).$$

20 Referring now to Fig. 2, we know that for elastic traffic, the graph of $B(S)$ against the service capacity S will be qualitatively similar to the graph drawn. The precise form of the function $B(S)$ is not known a priori. An essential feature of this invention is that it operates successfully without this knowledge. Another essential feature is the specific point S^* in Fig. 2, at which the $B(S)$ is equal to the service capacity for that
25 particular traffic.

It follows that, for an elastic source, the Optimum Service Rate is the unique value S^* for which $S=B(S)$. Hence any method of solving iteratively the equation $S=B(S)$ yields a method for determining S^* . In this specification, methods for solving the
30 equation iteratively and hence determining S^* are described. As mentioned above, these methods do not rely on first determining the function $B(S)$.

It will be appreciated that to carry out the invention, it is necessary to choose a suitable type of descriptor or descriptor format for allowing the bandwidth requirement

- 11 -

to be determined. As stated already, there are many such descriptors (D) and the more accurate and appropriate the descriptor (D) chosen or predetermined, the better will be the calculation of the BWR.

- 5 Assume the buffer size b to be known.
Fix the QoS target Q .
Configure the service rate at some initial value S_1 ,
Estimate the traffic descriptor for the source and use it together with the values b and Q to determine the BWR value $B_1 = B(S_1)$ see Fig. 3.
- 10 Next configure the service rate to be S_2 given by $S_2 = B_1$ see Fig. 4.

Re-estimating the traffic descriptor and repeating the procedure, we use the scheme:

- For $N = 1, 2, \dots$
- 15 Put $S_{N+1} = B_N$

- to get a sequence S_1, S_2, \dots
- As N increases, S_N converges to S^* , which is the unique value of the service capacity such that

- $S^* = B(S^*)$.
- It will be appreciated, by considering Fig. 2, that if $S_1 < B(S_1)$, the sequence increases to S^* ; if $S_1 > B(S_1)$, the sequence decreases to S^* .

- In practice, the service rates S_1, S_2, \dots may be reconfigured at suitable time intervals such as for example 5 minutes. It also has to be appreciated that in many instances it will be necessary to carry out the operation indefinitely. A reason for this would be for example where the software packages or operations being run are changed. In which case, the bandwidth requirement function changes to, for example, $B'(S)$. Now the operation has to be repeated and the new optimum service rate has to be obtained, in the sense that one now has to solve the new equation which is:

- 12 -

$$S = B'(S)$$

Referring now to the flow process of Fig. 5 which illustrates the iterative process of calculating the optimum service rate S^* . In step 10, the quality of service Q is set and the size of the buffer b is determined. A service rate S_1 is arbitrarily selected.
5 The traffic flowing through a buffer is measured by sampling the traffic. A traffic descriptor D is extracted from the sampled traffic which is representative of the statistical property flow of the traffic in step 11. In step 12, the BWR for that selected service rate, having a particular quality of service and a fixed buffer size, along with the traffic descriptor, provides an estimate of the required service capacity in real-time. In step
10 12, a new service rate is calculated from the estimated bandwidth requirement. In step 13, a check is made whether the new service rate coincides with the estimated bandwidth requirement. If they do not coincide, then steps 11 and 12 are repeated iteratively until the estimated bandwidth requirement coincides with the service rate which is the optimum service rate S^* . It will be appreciated that the extent to which
15 the BWR is required to coincide with the service capacity is determined by the network operator. In other words, the service rate S and the BWR do not have to exactly coincide, but be within a range that will maintain quality of service for the traffic flow. This process can be carried out at preset time intervals or alternatively when a change in the statistical property of the traffic flow is detected.

20

It will be appreciated that the invention can be carried out in many jurisdictions and that therefore it would not be unreasonable to find that certain operations are carried out in one country and more carried out another.

25

It is envisaged that there are many ways that the invention may be carried out by computers and similar equipment suitably programmed. Further the information and data generally supplied will allow many and varied control operations to be carried out using the invention.

30

Essentially what the present invention does is provide packet level quality of service on a data network where the rate at which a source emits data packets adapts to utilise all the available bandwidth. The present invention determines accurately the minimum service rate which is obviously the optimum service rate required to achieve a target level of service. For example in the remote querying of a database, the

application level response time increases with an increase in packet delay caused by the queuing of packets in the buffer at a network element. As has been explained above, in order to achieve acceptable response times ways must be found to achieve a quality of service target expressed as delay constraints. Present systems do not
5 allow this.

The present invention provides a way of directly controlling packet level quality of service. Essentially the present invention is "lightweight" and capable of being incorporated in closed loop control systems as described with reference to Fig. 4 will
10 allow the automation of the operation and avoids manpower intensive operations now required in changing the provisioning in a system with changing traffic patterns.

As will be appreciated by any mathematician, there are many iterative schemes to calculate the optimum service rate. Most of these have a trade off between speed of
15 conversion and computational overhead.

One of the great advantages of the present invention is that there is no need to know the actual function, but simply to carry out the calculations. Depending on the number of calculations that have to be carried out before convergence is reached, it
20 will be appreciated that if the results were to be plotted, much more accurate knowledge would be known regarding $B(S)$ than was known. However, this is not necessary and does not provide any great advantage to those operating the communications system.

25 Referring now to Fig. 6, there is illustrated part of a communications network and hardware configuration for carrying out the present invention, indicated generally by the reference numeral 30. A typical router or switch 31 receives a number of links 32 for the delivery and reception of data. Each link is connected to at least one buffer 33 incorporated in the switch 31. Each buffer 33 has an associated scheduler which schedules the service rate S for each buffer. In other words, the scheduler decides for each buffer 33 the maximum rate, measured in bits per seconds, that packets of
30 data can be forwarded or received by each buffer 33. Each buffer 33 is connected to a port 34 which in turn is connected to a larger link 35 which can receive or deliver data, depending on the application. This configuration is well known to the person

skilled in the area of communication switching systems. Each buffer is allocated a particular service rate by the scheduler. The service rate is controlled by providing a closed loop control system by having a programmable controller 36 connected to the router. The controller 36 sets the service rate for each buffer 33 incorporated in the
5 switch 31.

In operation, the closed loop control samples the data traffic flowing through the switch 31 at any particular time. The controller 36 comprises a database 37 and a computer 38. It will be appreciated that the database 37, for example a server, and
10 computer 38 can remotely operate with each other. They do not have to be in the same jurisdiction as the switch or indeed each other. The controller 36 selects an arbitrary service rate. This can be done by a user inputting a particular service rate via the computer 38 or can be selected by the database 37. The controller 36 extracts a traffic descriptor D which is a description of the statistical properties of the
15 traffic flow flowing through the buffer 33 at the time of the sampled traffic. The controller 36 calculates the BWR from the traffic descriptor D, the buffer size b and the required Quality of Service (QoS) for the buffer 33. The controller 36 estimates a new service rate from the calculated BWR. This is illustrated in Fig. 3 in which the calculated bandwidth requirement B_1 provides a new service rate S_2 . The controller
20 36 iteratively carries out the above steps until the estimated bandwidth requirement and the configured service rate coincide to provide a final service rate.

This final service rate is the optimum service rate at which the buffer 33 should be served. It will be appreciated that the computer and the database can be at separate
25 locations or indeed in separate jurisdictions, or alternatively housed on the one server, depending on the application, as long as closed loop control is achieved.

It will be appreciated that various aspects of the invention may be embodied on a computer that is running a program or program segments originating from a computer
30 readable or usable medium, such medium including but not limited to magnetic storage media (e.g. ROMs, floppy disks, hard disks, etc.), optically readable media (e.g. CD-ROMs, DVDs, etc.) and carrier waves (e.g., transmissions over the internet). A functional program, code and code segments, used to implement the present invention can be derived by a skilled computer programmer from the description of

- 15 -

the invention contained herein.

It will be appreciated therefore that a computerised program may be provided providing program instructions which, when loading into a computer, will constitute 5 the means for organising and rearranging the traffic flow in accordance with the invention and that this computer program may be embodied on a record medium, a computer memory, a read only memory or carried on an electrical carrier signal.

In this specification the terms "comprise, comprises, comprised and comprising" and 10 any variation thereof and the terms "include, includes, included and including" and any variation thereof are deemed to be totally interchangeable and should be afforded the widest interpretation possible.

This invention is in no way limited to the embodiment shown and may be varied in 15 both construction and detail within the scope of the claims.